DOCUMENT RESUME

ED 277 218 FL 016 065

AUTHOR Liskin-Gasparro, Judith E.
TITLE Teaching and Testing Oral Skills. ACTFL Master Lecture Series.
INSTITUTION Defense Language Inst., Washington, D.C.
PUB DATE Jun 83
NOTE 27p.; Paper presented by a member of the American Council on the Teaching of Foreign Languages (ACTFL) at the Defense Language Institute Foreign Language Center (Monterey, CA, June 1983).
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Achievement Tests; *Communicative Competence (Languages); *Curriculum Development; Interviews; Language Proficiency; Language Skills; *Language Tests; *Oral Language; *Rating Scales; Second Language Instruction; *Test Construction; Test Format; Test Reliability; Test Validity; *Verbal Tests

ABSTRACT

A discussion of teaching and testing oral skills in a second language looks at issues in the testing of speaking ability in general, the effect that testing can have on teaching, and the kinds of communicative activities that can be built into a foreign language program with oral proficiency as a goal. The issues addressed include the differential testing of the four language skills (listening, speaking, reading, and writing), reasons for testing, achievement versus proficiency tests, characteristics of good tests, the oral proficiency interview, theoretical and practical considerations in designing performance-based oral achievement tests, scoring oral classroom tests, and important concepts in the relationship of oral teaching and testing. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277218

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Liskin-Gasparro

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ACTFL MASTER LECTURE SERIES

Teaching and Testing Oral Skills

Judith E. Liskin-Gasparro

June 1983

Faculty and Staff Development Division
Assistant Dean for Instruction

DEFENSE LANGUAGE INSTITUTE
FOREIGN LANGUAGE CENTER
Presidio of Monterey, California 93944-5006

FOREWARD

For the past several years, prominent members of the American Council on the Teaching of Foreign Languages (ACTFL) have been presenting lectures to the faculty and staff of the Defense Language Institute, Foreign Language Center. The purpose of these lectures has been to discuss recent trends and developments in foreign language learning and teaching as well as to strengthen professional contacts between DLIFLC and ACTFL.

The ACTFL Master Lecture, "Teaching and Testing Oral Skills," by Dr. Judith E. Liskin-Gasparro, was presented at the DLIFLC in June 1983. This paper is published to make the content of the lecture fully accessible to the DLIFLC professionals.

The ideas and opinions expressed in this paper are those of the author and do not necessarily represent an official position of the DLIFLC nor of any other element of the United States Department of Defense.

Inquiries should be addressed to:

Commandant
Defense Language Institute
Foreign Language Center
ATTN: ATFL-DIN-FS
Dr. Lidia Woytak, Editor
Presidio of Monterey, CA 93944-5006

INTRODUCTION

The topic of teaching and testing oral skills in the classroom is one that has aroused a great deal of interest in foreign language professional circles in recent years. The DLI/FLC and other government language schools have perhaps been more aware of the importance of this area than have civilian academic institutions because the outcomes of government language training have been more precisely specified. Academic foreign language teachers have not traditionally been held accountable for the linguistic performance levels of their students, although this may now be changing.

This paper will discuss the testing of speaking ability in general, the effect that testing can have on teaching and, finally, the kinds of communicative activities that can be built into a foreign language program that has the development of oral proficiency as one of its primary goals.

TESTING THE FOUR SKILLS

The four language skills are usually grouped in a particular way for the purpose of teaching: listening and speaking -- the oral skills, and reading and writing -- the literacy skills. There are classroom activities that lend themselves well to each. For the purpose of testing, however, it makes sense to organize the skills differently -- skills of reception (listening and reading) on the one hand, and skills of production (writing and speaking) on the other.

One of the most important questions to consider in test construction is the choice of stimulus material. In tests of receptive skill the stimulus is always something that is read or heard. In a test of listening comprehension, for example, students might hear a tape, a human person, a record - but the stimulus is <u>heard</u>. Students indicate their responses in any one of a number of ways -- they can speak, mark a box, write something, point to a picture. The unifying element is that in tests of receptive skill it is the <u>stimulus</u> material that is

presented in the modality that is being tested.

In tests of productive skill, just the opposite is true. In a speaking test, for example, the stimulus can be almost anything -- something the students see (a picture), something they read, or something they hear. The common element in every case is that they have to respond by speaking, and the test consists in assessing the acceptability of the spoken response.

Performance Tests

Tests of speaking and writing are examples of performance tests, because the student has to produce language, and is then rated on that production.

Like any peformance test, an oral test can be either a discrete-point test or an integrative test, and can have a number of different formats. A discrete-point test is one that measures, one question at a time, mastery of small bits and pieces of language. It may, for example, include questions that ask the student to give the names of objects in a picture or to give a particular inflected verb or noun form, or to pronounce two words that constitute a minimal pair (caro-carro in Spanish, or boat and coat in English). An integrative test measures global language ability, i.e. the degree to which a student can put the bits and pieces of language together to perform a particular communicative function, like explaining how to reach the Presidio from the Cypress Tree Inn, or buying a train ticket to go from Frankfurt to Paris.

The notion of integrative performance tests is relatively new to foreign language education, but it exists as almost a given in other fields in which a skill is required to do the job or accomplish a particular task. Car mechanics, airplane pilots, hairdressers all have to demonstrate proficiency in their fields by a test of performance in order to be certified or licensed to do their jobs.

In the recent past, the MLA Cooperative Tests, which were widely used in the academic sector in the 1960s and 1970s, were the only

standardized tests available that tested speaking. The oral section was a discrete-point test. Students were asked to do such things as identify items in a picture and describe with a very short amount of speech what was happening in a particular situation. Students were scored on the basis of grammatical accuracy and appropriateness of vocabulary; the most common student strategy was to say as little as possible to avoid the possiblity of making errors.

The designers of the MLA tests opted for this discrete-point orientation out of concern for reliability (consistency) in scoring. Clearly, it is more possible to be consistent if the amount of speech to be evaluated is very limited. The trade-off, however, is a serious one; what was tested was not primarily language as communication.

The challenge for test development in the oral skills is to compose integrative tests of speaking ability. These are tests that, without a great deal of intervention from the tester, can measure students' ability to pull together what they have learned over time and to use it to produce sustained speech.

WHAT MAKES A GOOD TEST?

Test construction begins with a series of questions that the test developer asks himself or herself. The answers to those questions will determine the kind of test that is produced. The first decision is why we need the test in the first place. What are we trying to find out? More than any other factor, the answer to this question will determine the kind of testing instrument we need.

The first thing to decide, then, is <u>why</u> and <u>what</u> to test. There are several common reasons for testing. These can be divided into testing needs that are internal to the language program, and external needs, which serve the student and the student's future employer.

Internal Reasons for Testing

One common internal reason for testing is for placement. For example, a student arrives at DLI/FLC and says, "I've have three years of high school Russian. I liked Russian a lot, and I got straight A's." Where do you put him? You don't know the school, the teacher, or the textbook. Even if you did, you would still need a placement test to find out how the student's previous language training articulates with your program.

A second internal reason for testing is to assign grades. Even though the instructor may be very familiar with his or her students and know how they perform, the interests of fairness require that the instructor provide at various points during the course a common yardstick, a measurement that will be the same for all students. Students can then be asked to perform the same linguistic tasks under the same circumstances, and the instructor can see how they perform with respect both to the assigned tasks and to each other.

A third reason for testing is that testing provides motivation. Students are very practical creatures. Although they may have a certain commitment to language study, we can be sure that they are not in the classroom on time every day because the great passion of their lives is their language course. They are in class because they have a job to do, and they interpret that job as doing well on exams and getting good grades. As teachers, we can use this attitude to our advantage by making sure that we test them on what we want them to learn. If the students know, for example, that oral tests will be given regularly and that results will count significantly toward their course grade, then they are likely to take oral work in class and out of class more seriously.

A fourth reason for testing is to provide diagnostic feedback. When we give a test, we can discover students' areas of strength and weakness, where they have learned the material and where we as teachers need to spend more time. We can also get a sense of the effectiveness of our program. For example, if we spend a number of class hours on a

particular portion of the curriculum and the unit test reveals that the majority of students still have not mastered it to the desired degree, that is a signal to us to develop new teaching strategies for this particular section.

Achievement Tests and Proficiency Tests

All of the reasons for testing given above are internal to the language program, and they directly reflect the curriculum that has been set for a particular course. Most tests that correspond to these reasons for testing are achievement tests, which measure the degree to which students achieve the outcomes of a unit or a course or a program.

The distinction between achievement tests and proficiency tests is important in this context. Achievement tests are tied to a particular curriculum, and it is wise to use achievement tests throughout a language program for all of the reasons outlined above.

Achievement tests are usually scored with letter grades or number or percentage correct. These grades or scores can be very useful to the instructor in ranking students with respect to each other or in assessing their progress, but they are not terribly useful in communicating to a prospective employee about a student's ability. A proficiency test would be better for outside use. If a company wants to hire a receptionist who can answer the phone, greet visitors, and arrange appointments in German, the personnel department will not be interested in how many language courses the applicant has taken. They will want to know how well the individual can function in the language on the job. For this a proficiency test is needed, one that will be curriculum-free, and whose rating will be based on criteria of real-life language use.

Any language program needs both kinds of tests: achievement tests during the course, and a proficiency test at the end to assess the students' skills compared to those of educated native speakers.

CHARACTERISTICS OF GOOD TESTS

After the important first decision has been made about the purpose of the test, one can begin to construct an appropriate testing instrument. All good tests will share some characteristics:

(1) Validity. A valid test is one that actually measures what it purports to measure. One example of a non-valid test of speaking ability is a test in which students read aloud, and for which grades are given on the basis of pronunciation and fluency. Such a test may be a very valid measure of pronunciation, but it is not a valid test of speaking, since students are not asked to speak.

Sampling is an important component of test validity. Intensive language courses, such as those at DLI, meet for several hours each day and cover a great deal of material in a relatively short period of time. If the instructor wishes to give a test on a four-week unit, he or she cannot test everything that has been taught. The instructor has to chose topics and skills that are representative of that four-week period.

(2) Reliability. A second feature of all good tests is reliability. Reliability in this context has to do with the consistency of the measure. Let's assume that we have a valid test, one with well thought-out questions at a level that will be accessible to students. If we give the test today and tomorrow to the same group of students (and we eliminate the possibility of advantage due to familiarization with the test), a reliable test will give us equivalent results. In order for this to happen, the administration of the test has to be carefully replicated and, most important, the standards used in scoring have to be consistent from one test administration to the next. This becomes especially critical when we are dealing with performance tests, which are scored by people, who are less consistent than machines.

(3) A third feature of successful tests is emphasis on the important points. Most of us in the academic community operate with text materials that follow a grammatical syllabus. When a particular concept or structure is taught, all of the possibilities and exceptions are presented. A good test will not require students, especially at the beginning of language training, to demonstrate that they have learned everything that they have been exposed to. It is far better to construct a test that measures the main points or concepts.

(4) The last feature of good tests is one that is implied in everything mentioned thus far. A good test is a microcosm of the activities and exercises and materials of the course whose objectives it is measuring. When students complain that a test is "not fair", what they are most often saying is that they were confronted with something terribly unfamiliar. If students have never written a composition in class, it is not fair to make such a task part of the final examination. If they have never been asked to role play in class, it is unfair to ask them to do that for the first time on the examination.

TESTING AND TEACHING

The features of good tests described above carry with them an underlying principle: successful testing and succesful teaching go hand in hand. As Dean Ray Clifford has said, "Students learn to do what they practice doing." This statement reflects the importance of articulating course outcomes, curriculum, teaching strategies, and testing.

Students traditionally try to get the best results from the least effort. In literature classes they read plot summaries instead of the novels or plays; in science classes they learn equations and formulas, and in foreign language classes they try to do well on tests by stuffing their heads full of vocabulary and grammar rules the night before a test. If we want them to integrate what they learn and speak the language in free conversation, then we have to teach them to do that in class and test them on those same skills in our tests.

The Oral Proficiency Interview

As explained above, proficiency tests are appropriate instruments to measure students' ability at the end of a sequence of courses or a major linguistic experience, such as study abroad. The oral proficiency test is perhaps the most successful proficiency test in our field. It follows the four features of good tests discussed above.

(1) Validity. The oral interview is a test of speaking, and indeed test candidates are required to speak the language during the interview.

(2) Reliability The 0-5 scale and the procedures for interviewing to get a ratable sample are constant from interview to interview, even though the topics and the questions may vary. What makes the rating reliable is the training that is given to the testers, which teaches them to apply the same standards to each interview. An important part of this process is the 30 years of government and, more recently, academic experience with the rating scale, which allow us to describe accurately what kinds of language are characteristic of each level.

(3) Test the important parts. In an oral proficiency interview, the domain that can be tested is the entire language as it is used by native speakers. But testers start from the bottom up, sampling from the body of the language only to the point where the examinee's language is no longer adequate to the task. There are no discrete-point questions of grammar or vocabulary; what is tested is the student's ability to handle major linguistic functions and content areas.

(4) Microcosm. The oral proficiency interview is indeed a microcosm of all possible conversations within linguistic reach of the candidate. An important part of the art of inteviewing

is the ability to elicit a representative sample of candiate speech in a relatively short period of time.

The oral proficiency interview works very well as an instrument for summative evaluation, to find out what students are able to do after a course of study. As a type of formative evaluation, however, it will most often not be very helpful. The ranges 0, 1, 2, 3, etc. even with plusses, are very broad, and the categories do not provide a great deal of discrimination power. If we give a student an oral proficiency interview on a Friday, for example, and then again the next Friday after 30 additional hours of instruction, the chances are that the student will get the same rating. The rating will not give the teacher a good sense of the progress the student has made. Similarly, the interview will not be a good testing instrument to differentiate among students in the same class for grading purposes. The chances are they will all get the same rating, or be within a plus of each other.

For formative evaluation purposes, then, we want to devise speaking tests that are more closely tied to the curriculum. They should be achievement-oriented, but nevertheless integrative and communicatively based, tests that reflect a classroom in which students are given opportunities to interact verbally with the instructor and with each other with minimal teacher intervention.

DESIGNING PERFORMANCE-BASED ORAL ACHIEVEMENT TESTS

Although performance testing is a somewhat revolutionary concept in the academic world, it is a commonplace in the vocational field, where candidates for technical or professional certification must demonstrate their ability to carry out the tasks involved in their future jobs. Ryans and Frederiksen[1] have devised the following list of stages that are needed for the development of performance tests.

DEVELOPMENT OF PERFORMANCE TESTS

(1) Job analysis
(2) Select tasks that represent the job
(3) Develop rating form
(4) Practical considerations
(5) Pretest
    o revise as necessary
(6) Directions for use
(7) Administer and score performance test

Table 1

These steps can be applied to the development of oral tests as well. The considerations and tasks that accompany each step are as follows:

(1) Job analysis

What is the linguistic and/or social task that we want to test? To accomplish that task, how is language used? With whom must the speaker communicate, and about what?

(2) Select tasks

This is the construction of questions or content for the oral test.

(3) Develop rating form

How will we grade this test? The development of the rating scale and the criteria for assigning students to a place on the scale will determine the reliability of the scoring of the speaking test. Further discussion of the grading of oral tests follow below.

(4) Practical considerations

Although they may seem insignificant in principle, practical considerations can make or break a testing program. For example, if

students are known to be at their worst on Monday morning or Friday afternoon, it would not be wise to schedule oral tests for those times. If a test requires language lab facilities, the instructor should make sure that they are available when they are needed. If students are to be tested individually by the instructor, the instructional needs of the rest of the class must be met while the instructor is occupied with the testing. If students have the opporunity to communicate with each other before they have all taken the test, then the format of the test will have to be one that does not give an advantage to the last students who take it.

(5) Pretest

The purpose of pretesting is to observe how a test or test questions function, and then to make revisions as appropriate. The pretest, then, is a test of the test itself, and of the instructor as test maker.

The importance of pretesting cannot be over-estimated. Even the best-conceived testing programs will have some flaws. It is important, therefore, to try out new testing ideas and procedures in a setting in which the instructor does not have to depend on the results for grading purposes. It can also be very helpful to include the students in the pretesting process by polling them about their reactions to test format, content, and level of difficulty.

(6) Directions for use

The wording of the directions should be clear and straightforward, so that students understand exactly what is expected of them. Based on a philosophy born in the era of contract learning and self-paced instructions, many teachers now tell students ahead of time how they will be graded in a course and in written work, such as exams and compositions. This is a very positive development. By removing some of the mystery from the evaluation process, instructors enable students to assume greater responsibility for their learning.

Clear and complete directions for the test administrator are also an important factor in successful testing. The overall goal is to ensure consistency, and therefore fairness, to all students.

SCORING ORAL CLASSROOM TESTS

Classroom tests have traditionally been decidedly discrete-point in orientation; that is, they test small bits of information one at a time. A performance test by definition measures global and integrative language skills. When grading classroom tests, teachers have to find the happy medium between these two extreme styles of testing. The grading system should include both an assessment of the overall communication, as well as measurement of the accuracy of the constituent parts -- fluency, grammar, pronunciation, comprehensibility, etc. Most teachers find it difficult to evaluate so many linguistic and communicative factors at the same time, and will understandably slip unconsciously into a grading system that emphasizes one aspect of language skill over all of the others.

It is important to realize that one cannot make as fine distinctions in an oral performance test as one can with a discrete-point test. Simplifying a complex measurement concept, there is an inverse relationship between the number of points on the scoring scale and the reliability of the scoring. For example, if an instructor tries to discriminate very finely among students by developing a scoring scale of 30 or 40 points, the instructor will be unable to maintain consistent standards from one student to the next, and the reliability of the test will suffer.

The scoring technique proposed below uses not a single scale but a combination of scales that covers the several linguistic and communicative factors involved in oral communication. By separating oral communication into its component parts for purposes of evaluation while still measuring the overall communicative effect, the instructor can move toward the "happy medium" of evaluation described above. The scales allow the instructor to focus on one or more aspects of oral language and to keep track of all of them at the same time.

The scales are adapted from rating scales developed by Walter Bartz in a book entitled Testing Oral Communication in the Classroom.[2] The first change from Bartz's system consists in reducing the number of points on the scales to provide for more reliable measurement. The scales now contain four points, along with a point 0 at the bottom that would hardly ever be used. The scond revision concerns the description of each point on the scale. With all of the scales, it is important to remember that the instructor will be grading the students within a restricted range. The 0-4 scale is not in any way related to the 0-5 oral proficiency scale. The students may all be at Level 1 on the oral proficiency scale, and it is for this very reason that an oral proficiency interview is not an appropriate testing instrument for formative, mid-course evaluation. The 0-4 scale is based on realistic expectations that the instructor has for students at a particular point in their training.

Let us take as an example the evaluation of fluency. Fluency is defined not as speed of delivery, but as the overall smoothness and naturalness of the language. The 0 point here is reserved for performances that are virtually no language at all. The points 1-4 are then divided into an lower half (1-2) and an upper half (3-4). Students in the lower half are performing up to about one-half of the instructor's expectations, and students in the upper half are doing close to or as much as the instructor can reasonably expect of students at a particular point in their training. The fluency scale, then, is as follows:

FLUENCY

Overall smoothness and naturalness of speech

0	So halting as to be virtually silence
1	Very halting; fragmentary delivery
2	Frequent halting, unnatural pauses
3	Few unnatural pauses, fairly smooth delivery
4	Smooth, natural delivery

Table 2

The same procedure can be followed in developing rating scales for other aspects of oral performance, such as comprehensibility, amount of communication, grammatical correctness, and effort to communicate. Scales for these linguistic factors can be found in the Appendix.

Instructors and students may find the "Amount of Communication" scale particularly useful as a motivational device. Students tend to give short responses, in the hope of being grammatically correct. As a result, they may become low linguistic risk-takers, which in the long run may retard progress in language learning. If students know that they will be rewarded in an oral test for the amount they speak, they will push themselves to say more than the absolute minimum. A word of caution is in order, however. There is a very delicate balance between the desire to communicate and linguistic accuracy, and perhaps the instructor's most important task will be to guide students between these two poles of language production.[3]

Tips for Using the Rating Scales

As with any new endeavor, it is important to move slowly and thoughtfully when trying out a different testing system. The following suggestions are offered to those who may wish to adopt a holistic method of grading oral tests.

(1) Don't try to use all of the scales at once, especially at the beginning. It would be advisable to start by using and becoming familiar with one of the scales. A good one to start with might be the "Amount of Communication" scale, since it may represent a change in orientation for both students and instructors. After the instructor has learned how to rate amount of communication accurately and consistently, he or she can add another scale, and so on.

(2) Instructors will probably find that they want to develop a scoring sheet to use in grading each student. The advantage of a scoring sheet is that it provides data to use in discussing test results

with an individual student, or in looking for patterns of strength and weakness in a group of students.

(3) If testing time is a problem, the instructor can learn, with experience, to score a goup of students at the same time. The students can be given a topic to discuss, or a linguistic task to carry out in pairs or in a small group. The instructor can observe the communication, and grade students as they speak in the interactive situation.

(4) It is important for the instructor to remember that his or her expectations for a 1, 2, 3, or 4 will change as the class progresses. Clearly, the instructor's expectations for students after four weeks will be very different from the expectations after 16 weeks or 20 weeks. The scale points represent relative performances, not fixed criteria as they do in the oral proficiency interview.

(5) Work with your colleagues in the setting of standards for student performance at a particular point in the language program. This kind of collaboration not only is beneficial in test development work, but also may initiate a process of cooperation that can improve the program overall.

(6) Expect to have to experiment. As instructors gain experience with performance-based oral testing, they will want to change and refine the rating scales. This is a natural and valuable evolutionary process that should be approached with enthusiasm.

TOUCHSTONES OF ORAL TEACHING AND TESTING

One very interesting insight that one gains by developing oral performance tests in a program devoted to functional language ability is that the classroom activities and testing formats will be in many cases the same. This allows students to practice in class the very skills that are tested, and that the instructor wants them to learn.

Personalization is a technique that applies equally well to teaching and testing. Students are far more interested in talking about themselves than about non-existent characters invented by the textbook author. This intrinsic interest on the part of the students will motivate them to communicate facts about their lives and to find out similar information from their peers.

Contextualization is a principle that only recently has come to the fore. Most traditional text materials treat language as though it were a series of non-sequiturs, as though one sentence were hardly ever related to the next one. Sentences for translation or for practice of a particular structure are most often thematically unrelated to each other. The result is that students treat language at the sentence level only, and have a great deal of difficulty joining sentences together into paragraphs. Contextualization in the classroom consists of building exercises around a common semantic theme, as well as a common grammatical theme. In this way, students see how sentences relate to each other and perceive language as an integrated whole.

Another important part of teaching and testing oral skills is the development of a student-oriented classroom. Students in teacher-centered classrooms have not learned to depend on their own lingusitic resources; they constantly look to the teacher for confirmation or correction. In a student-centered setting, in which the teacher puts himself or herself in the background and has the students interact with each other, the students expand their communicative resources and become more linguistically independent.

Teachers who learn about the oral proficiency interview discover that it confirms a belief that they have long held, but perhaps did not know how to articulate, i.e. that there is a hierarchy of linguistic functions. Students acquire first the ability to list and name single words and phrases. Then they can deal with simple sentences, and only after that can they join sentences together into paragraphs to narrate or describe. Finally, they learn to discuss abstract topics, support opinions, and hypothesize. The recognition that these functions are hierarchical has important implications for curriculum design and the development of classroom activities.

Finally, it is important in the teaching and testing of oral skills to take a positive approach. In multiple-choice testing, students are often able to hide what they do not know. In order to select the right answer, they need to know only enough to pick out the right choice and/or eliminate the incorrect choices. On the other hand, when we put students in a position of performing, either in the classroom or on a test, they are vulnerable. Because of the nature of performance testing, students reveal not only their strengths, but also their errors and imperfections. It is important, therefore, to remember to focus more on the positive aspects of the performance than on the errors.

CONCLUSION

In conclusion, perhaps the key to successful performance test development and to performance-oriented teaching is the willingness to experiment. One of the fundamental truths about teaching is that there is no single "right" way to do things. Any method, technique, or test is more or less successful depending on a variety of factors--students' personalities, the instructor's preferences, and the ability level of the students. An approach that works well one year may be a complete disaster the next. The key to success is to bring our experience and knowledge freshly each day to the task of helping students acquire functional language ability.

NOTES

[1]David G. Ryans and Norman Fredericksen, "Performance Tests of Educational Achievement," in Educational Measurement, ed. E. F. Lindquist (Washington, DC: American Council on Education, 1951), pp. 455-94.

[2]Walter H. Bartz, Testing Oral Communication in the Foreign Language Classroom, in Language in Education: Theory and Practice, No. 17 (Washington, DC: Center for Applied Linguistics, 1979).

[3]See Theodore V. Higgs and Ray T. Clifford, "The Push Toward Communication," Curriculum, Competence, and the Foreign Langauge Teacher, ed. Theodore V. Higgs, ACTFL Foreign Language Education Series, Volume 13 (Lincolnwood, IL: National Textbook Company, 1982), pp. 57-79 for further discussion of this subject.

APPENDIX

Sample Scales for Scoring Oral Tests

COMPREHENSIBILITY

0 No comprehension at all by the teacher

1 Teacher comprehended only isolated words, phrases

2 Teacher comprehended about half

3 Teacher comprehended most, but not all

4 Complete comprehension by the teacher

QUALITY (GRAMMATICAL CORRECTNESS) OF COMMUNICATION

0 No utterance rendered correctly

1 Very few utterances correct

2 Some (up to half) utterances correct

3 Many utterances correct, but some problems with structure remain

4 All or most utterances correct; errors are either very minor or concern difficult structures that student attempts but has not yet learned

EFFORT TO COMMUNICATE/CREATIVITY

Willingness to "risk" linguistically to get message across; attempts circumlocution and paraphrase; tries several ways to say something if the first repetition or clarification to advance conversation

0 None

1 Very little (lots of embarrassed silence)

2. Some effort

3 Considerable effort

4 Extraordinary effort and creativity

AMOUNT OF COMMUNICATION

Quantity of information related to the communicative situation or task

0 None

1 Very little

2 Some (about half of what was expected)

3 Most

4 All